A COUNTEREXAMPLE TO THE POSSIBILITY OF AN EXTENSION OF THE ECKART-YOUNG LOW-RANK APPROXIMATION THEOREM FOR THE ORTHOGONAL RANK TENSOR DECOMPOSITION*

TAMARA G. KOLDA†

Abstract. Earlier work has shown that no extension of the Eckart-Young SVD approximation theorem can be made to the strong orthogonal rank tensor decomposition. Here, we present a counterexample to the extension of the Eckart-Young SVD approximation theorem to the orthogonal rank tensor decomposition, answering an open question previously posed by Kolda [*SIMAX*, 23(1):243–355, July 2001].

Key words. singular value decomposition, principal components analysis, multidimensional arrays, higher-order tensor, multilinear algebra

AMS subject classifications. 15A69, 49M27, 62H25

1. Introduction.

We consider the problem of whether or not we can extend the Eckart-Young result to tensors for a particular extension of the SVD known as the *orthogonal rank decomposition*. In other words, suppose a tensor A has an orthogonal rank decomposition of the form

$$A = \sum_{i=1}^{r} \sigma_i U_i.$$

Here, r is the minimal number of terms that can be used to represent A, the σ_i's are scalars such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and the U_i's are *decomposed tensors* (i.e., rank-1 tensors) with the property that any pair of the decomposed tensors are *orthogonal*. Notation and definitions are provided in §2. The question is: Does the sum of the first k terms yield the best rank-k approximation?

In the case that A is a matrix, the orthogonal rank approximation is equivalent to the SVD approximation where each σ_i is equal to the ith singular value and each U_i is the outer product of the ith left singular vector with the ith right singular vector. For matrices, the Eckart-Young theorem [3] says that the best rank-k approximation to A is indeed given by the sum of the first k terms of the SVD.

Kolda [4] showed that the Eckart-Young approximation property does not hold for the strong orthogonal rank tensor decomposition, another extension of the SVD. Leibovici and Sabatier have attempted to show that the Eckart-Young approximation property holds for the orthogonal rank tensor decomposition [5, Theorem 2]. The refutation of that claim in [4] is incorrect,[1] so here we reconsider this issue and show that the Eckart-Young approximation property does not hold for the orthogonal rank tensor decomposition.

Our argument proceeds as follows. In §3, we present an orthogonal rank decomposition of a tensor A. From the decomposition, we can determine that the orthogonal

*This work was supported by the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy, under contracts DE–AC05–96OR22464 with Lockheed Martin Energy Research Corporation and DE-AC04-94AL85000 with Sandia Corporation.

†Computational Science and Mathematics Research Department, Sandia National Laboratories, Livermore, CA 94551–9217, tgkolda@sandia.gov.

[1] It was also erroneous to refer to Remark 6.3 of [5] as a "result" rather than a "remark".

rank of A is 2. If the Eckart-Young extension hypothesis is true, then the first term of the decomposition should be the best rank-1 approximation of A. In §4, however, we compute the best rank-1 approximation of A and find that it is not equal to the first term of the orthogonal decomposition presented in §3. We know from Kolda [4] that the orthogonal rank decomposition is not unique, so in §5, we consider whether or not we can extend the best rank-1 approximation of A to an orthogonal rank decomposition. We find that the best we can possibly do is produce a 3-term orthogonal decomposition, which is not a rank decomposition. Thus we conclude in §6 that the Eckart-Young approximation theorem for the SVD cannot be extended to the orthogonal rank tensor decomposition.

2. Notation & Definitions. We use the notation and definitions from Kolda [4], briefly summarized here. If A is an $m_1 \times m_2 \times \cdots \times m_n$ tensor, we say the *order* of A is n, and the jth *dimension* of A is m_j. The set of all tensors of size $m_1 \times m_2 \times \cdots \times m_n$ is denoted by $\mathcal{T}(m_1, m_2, \ldots, m_n)$.

Decomposed tensors are the building blocks of tensor decompositions. A *decomposed tensor* is a tensor $U \in \mathcal{T}(m_1, m_2, \ldots, m_n)$ that can be written as

$$U = u^{(1)} \otimes u^{(2)} \otimes \cdots \otimes u^{(n)},$$

where $\otimes$ denotes the outer product and each $u^{(j)} \in \mathbb{R}^{m_j}$ for $j = 1, \ldots, n$. The vectors $u^{(j)}$ are called the *components* of U. The set of all decomposed tensors of size $m_1 \times m_2 \times \cdots \times m_n$ is denoted by $\mathcal{D}(m_1, m_2, \ldots, m_n)$.

Let $U, V \in \mathcal{D}(m_1, m_2, \ldots, m_n)$. We say that U and V are *orthogonal* ($U \perp V$) if

$$U \cdot V = \prod_{j=1}^{n} u^{(j)} \cdot v^{(j)} = 0.$$

The *orthogonal rank* of A, denoted $\operatorname{rank}_\perp(A)$, is defined to be the minimal r such that A can be expressed as

$$A = \sum_{i=1}^{r} \sigma_i U_i,$$

where $U_i \perp U_j$ for all $i \neq j$ and $\|U_i\| = 1$ for all i. This decomposition is called the *orthogonal rank decomposition*. Other decompositions are described by Kolda [4], including the *strong orthogonal rank decomposition* mentioned in §1.

3. An Example Tensor with Orthogonal Rank 2. Consider the following tensor $A \in \mathcal{T}(m, m, m)$ defined by

$$A = \sigma_1 \underbrace{a \otimes a \otimes a}_{U_1} + \sigma_2 \underbrace{b \otimes b \otimes \hat{a}}_{U_2}. \tag{3.1}$$

Let the vectors $a, \hat{a} \in \mathbb{R}^m$ be orthogonal (i.e., $a \perp \hat{a}$) with $\|a\| = \|\hat{a}\| = 1$. Define $b = \frac{1}{\sqrt{2}}(a + \hat{a})$. Let $\sigma_1, \sigma_2 \in \mathbb{R}$ with $\sigma_1 > \sigma_2 > 0$. Observe that $U_1 \perp U_2$, so $\operatorname{rank}_\perp(A) \leq 2$. Further, we can see that we cannot reduce this to a single decomposed tensor since the span in *every* component has dimension 2. Thus, we can conclude that

$$\operatorname{rank}_\perp(A) = 2$$

and that (3.1) is an orthogonal rank decomposition of A.

4. The Best Rank-1 Approximation. We directly compute the best rank-1 approximation of A in (3.1), which we denote by

$$A_1 = \gamma\, x \otimes y \otimes z, \tag{4.1}$$

where $\gamma > 0$ and $\|x\| = \|y\| = \|z\| = 1$. Note that we may assume that γ is positive since its sign can be absorbed into, e.g., the x-vector without affecting the quality of the approximation. We proceed to solve for γ, x, y, z.

Consider the first component. Without loss of generality, we assume $x \in \operatorname{span}\{a, \hat{a}\}$. Let $\hat{x}$ be the orthogonal complement of x in the space defined by $\operatorname{span}\{a, \hat{a}\}$. Then we can define $\alpha_x, \beta_x \in \mathbb{R}$ such that

$$\begin{aligned}
x &= \alpha_x a + \beta_x \hat{a},\\
\hat{x} &= \beta_x a - \alpha_x \hat{a},\\
a &= \alpha_x x + \beta_x \hat{x}, &(4.2)\\
\hat{a} &= \beta_x x - \alpha_x \hat{x}, &(4.3)
\end{aligned}$$

Using these definitions, we can express b as

$$b = \frac{(\alpha_x + \beta_x)}{\sqrt{2}}\, x - \frac{(\alpha_x - \beta_x)}{\sqrt{2}}\, \hat{x}$$

We can produce similar decompositions for the second and third components using y and z, respectively. We can then rewrite A in terms of x and $\hat{x}$ in the first component, y and $\hat{y}$ in the second component, and z and $\hat{z}$ in the third component; in other words, we can rewrite A as the sum of eight terms which are the combinations of $\{x, \hat{x}\} \otimes \{y, \hat{y}\} \otimes \{z, \hat{z}\}$ as follows:

$$\begin{array}{rcll}
A & = & [\, \sigma_1\, \alpha_x\alpha_y\alpha_z + \frac{\sigma_2}{2}\, (\alpha_x+\beta_x)\,(\alpha_y+\beta_y)\,\beta_z \,] & x \otimes y \otimes z\\
& + & [\, \sigma_1\, \alpha_x\alpha_y\beta_z - \frac{\sigma_2}{2}\, (\alpha_x+\beta_x)\,(\alpha_y+\beta_y)\,\alpha_z \,] & x \otimes y \otimes \hat{z}\\
& + & [\, \sigma_1\, \alpha_x\beta_y\alpha_z - \frac{\sigma_2}{2}\, (\alpha_x+\beta_x)\,(\alpha_y-\beta_y)\,\beta_z \,] & x \otimes \hat{y} \otimes z\\
& + & [\, \sigma_1\, \alpha_x\beta_y\beta_z + \frac{\sigma_2}{2}\, (\alpha_x+\beta_x)\,(\alpha_y-\beta_y)\,\alpha_z \,] & x \otimes \hat{y} \otimes \hat{z}\\
& + & [\, \sigma_1\, \beta_x\alpha_y\alpha_z - \frac{\sigma_2}{2}\, (\alpha_x-\beta_x)\,(\alpha_y+\beta_y)\,\beta_z \,] & \hat{x} \otimes y \otimes z\\
& + & [\, \sigma_1\, \beta_x\alpha_y\beta_z + \frac{\sigma_2}{2}\, (\alpha_x-\beta_x)\,(\alpha_y+\beta_y)\,\alpha_z \,] & \hat{x} \otimes y \otimes \hat{z}\\
& + & [\, \sigma_1\, \beta_x\beta_y\alpha_z + \frac{\sigma_2}{2}\, (\alpha_x-\beta_x)\,(\alpha_y-\beta_y)\,\beta_z \,] & \hat{x} \otimes \hat{y} \otimes z\\
& + & [\, \sigma_1\, \beta_x\beta_y\beta_z - \frac{\sigma_2}{2}\, (\alpha_x-\beta_x)\,(\alpha_y-\beta_y)\,\alpha_z \,] & \hat{x} \otimes \hat{y} \otimes \hat{z}.
\end{array} \tag{4.4}$$

The coefficient of the $x \otimes y \otimes z$ term is

$$\gamma = \sigma_1 \alpha_x \alpha_y \alpha_z + \frac{\sigma_2}{2}(\alpha_x + \beta_x)(\alpha_y + \beta_y)\beta_z.$$

The best rank-1 approximation of the form in (4.1) is produced by maximizing γ [2]:

$$\begin{array}{ll}
\max & \sigma_1 \alpha_x \alpha_y \alpha_z + \frac{\sigma_2}{2}(\alpha_x + \beta_x)(\alpha_y + \beta_y)\beta_z\\
\text{s.t.} & \alpha_x^2 + \beta_x^2 = 1\\
& \alpha_y^2 + \beta_y^2 = 1\\
& \alpha_z^2 + \beta_z^2 = 1
\end{array} \tag{4.5}$$

First observe that none of the α's can be zero because in that case we have

$$\gamma = \frac{\sigma_2}{2}(\alpha_x + \beta_x)(\alpha_y + \beta_y)\beta_z \le \frac{\sigma_2}{2}(\sqrt{2})(\sqrt{2})(1) = \sigma_2.$$

From the assumption that $\sigma_1 > \sigma_2$, we can get a larger objective value by simply choosing $\alpha_x = \alpha_y = \alpha_z = 1$ to yield $\gamma = \sigma_1$.

It also turns out that the β's are nonzero, but proving this is more difficult. We must consider the first-order necessary conditions for optimality for (4.5), which produces the following system of equations:

$$\sigma_1 \alpha_y \alpha_z + \frac{\sigma_2}{2}(\alpha_y + \beta_y)\beta_z + 2\lambda_x \alpha_x = 0 \tag{4.6}$$

$$\sigma_1 \alpha_x \alpha_z + \frac{\sigma_2}{2}(\alpha_x + \beta_x)\beta_z + 2\lambda_y \alpha_y = 0 \tag{4.7}$$

$$\sigma_1 \alpha_x \alpha_y + 2\lambda_z \alpha_z = 0 \tag{4.8}$$

$$\frac{\sigma_2}{2}(\alpha_y + \beta_y)\beta_z + 2\lambda_x \beta_x = 0 \tag{4.9}$$

$$\frac{\sigma_2}{2}(\alpha_x + \beta_x)\beta_z + 2\lambda_y \beta_y = 0 \tag{4.10}$$

$$\frac{\sigma_2}{2}(\alpha_x + \beta_x)(\alpha_y + \beta_y) + 2\lambda_z \beta_z = 0 \tag{4.11}$$

Case I. We show $\beta_z \neq 0$ by contradiction. Suppose $\beta_z = 0$. Note that this implies $\alpha_z = \pm 1$ from the equality constraint in (4.5). From (4.9) and (4.10), we get $\lambda_x \beta_x = 0$ and $\lambda_y \beta_y = 0$. Suppose $\lambda_x = 0$. Then we get that $\alpha_y = 0$ from (4.6), but we know none of the α's are zero from the argument above, so this is a contradiction and $\lambda_x \neq 0$. Likewise, we can show $\lambda_y \neq 0$. So, we must have $\beta_x = \beta_y = 0$ and $\alpha_x = \alpha_y = \pm 1$, but then (4.11) yields a contradiction. Thus we conclude that $\beta_z \neq 0$.

Case II. We show $\beta_x \neq 0$ by contradiction. Suppose $\beta_x = 0$. Then from (4.9), we have $(\alpha_y + \beta_y)\beta_z = 0$. From Case I, we know that $\beta_z \neq 0$, so we must have $(\alpha_y + \beta_y) = 0$. Combining this with (4.11) and the fact that $\beta_z \neq 0$, we get $\lambda_z = 0$. Then from (4.8), we get $\alpha_y = 0$ since $\alpha_x = \pm 1$. Once again, since none of the α's can be zero, we have a contradiction. Hence, we must have $\beta_x \neq 0$.

Case III. Using an argument analogous to Case II, we can show that $\beta_y \neq 0$.

Thus we have that every α and β is nonzero, i.e,

$$\alpha_x \neq 0,\ \alpha_y \neq 0,\ \alpha_z \neq 0,\ \beta_x \neq 0,\ \beta_y \neq 0,\ \text{and } \beta_z \neq 0. \tag{4.12}$$

This implies that each component of A_1, the best rank-1 contribution to A, has contributions from both a and $\hat{a}$. Therefore, $A_1 \neq U_1$; i.e., A_1 is not the first term of the orthogonal rank decomposition given in (3.1). In the next section, we attempt to extend A_1 to an orthogonal rank decomposition.

Before we go on, let us show that we may, without loss of generality, assume that all the α's and β's are positive. The argument is as follows.

At any optima of (4.5), each term of γ must be nonnegative. If the first term were negative, we could reverse the sign of α_z, which is nonzero by (4.12), resulting in a larger objective value without affecting the other term nor violating the constraint. Likewise for the second term and β_z. Thus,

$$\alpha_x \alpha_y \alpha_z > 0 \quad \text{and} \quad (\alpha_x + \beta_x)(\alpha_y + \beta_y)\beta_z > 0 \tag{4.13}$$

Additionally, for any optima of (4.5), we must have

$$\operatorname{sign}(\alpha_x) = \operatorname{sign}(\beta_x) \quad \text{and} \quad \operatorname{sign}(\alpha_y) = \operatorname{sign}(\beta_y). \tag{4.14}$$

In this case, if α_x is positive and β_x is negative or vice versa, reversing the sign of whichever one is not the same as their sum, $(\alpha_x + \beta_x)$, results in a larger objective value without affecting the other term nor violating the constraint. Note that here we assume that if the sum is negative, there is one other negative term in the product which enforces the positivity required by (4.13).

Finally, for any optima of (4.5), we must also have

$$\operatorname{sign}(\alpha_z) = \operatorname{sign}(\beta_z) \tag{4.15}$$

If α_x and α_y are both negative or both positive, then α_z must be positive to ensure that the first term of γ is positive from (4.13). Furthermore, this implies that $(\alpha_x+\beta_x)$ and $(\alpha_y + \beta_y)$ are both negative or both positive by (4.14), so once again β_z must be positive to ensure positivity of the second term of the objective. Likewise, both α_z and β_z must be negative if α_x and α_y have opposite signs.

From (4.14) and (4.15), each (α, β) pair must have the same sign. Now suppose that an (α, β) pair, say the one associated with x, is negative. Then we may *absorb* the minus sign by substituting $x = -x$ and $\hat{x} = -\hat{x}$ in (4.2) and (4.3). Therefore we may assume, without loss of generality, that

$$\alpha_x > 0,\ \alpha_y > 0,\ \alpha_z > 0,\ \beta_x > 0,\ \beta_y > 0,\ \text{and } \beta_z > 0. \tag{4.16}$$

5. Extending the Rank-1 Approximation. Although the best rank-1 approximation to A is not the first term of the orthogonal decomposition in §3, there is still the possibility that the best rank-1 approximation may be the first term of some *alternate* orthogonal rank decomposition of A since we know that the decomposition is not unique [4, Lemma 3.5]. Therefore we consider the problem of extending the best rank-1 approximation to an orthogonal rank decomposition, i.e., an orthogonal decomposition with only two terms.

Now consider the remainder tensor $R_1 = A - A_1$, consisting of last seven terms from (4.4). In order for us to extend the best rank-1 approximation defined by A_1 to an orthogonal decomposition of rank 2, we must be able to rewrite R_1 as a *single decomposed tensor* for any choice of σ_1 and σ_2.

From (4.16), we know that all of the α- and β-terms are positive. Observe that as the ratio $\sigma_1/\sigma_2 \to +\infty$, we have $\alpha_x, \alpha_y, \alpha_z \to 1$. In other words, there exists σ_1 sufficiently larger than σ_2, such that

$$\alpha_x > \beta_x \text{ and } \alpha_y > \beta_y. \tag{5.1}$$

If we choose σ_1 and σ_2 such that (5.1) holds, then the coefficients in R_1 corresponding to $x \otimes \hat{y} \otimes \hat{z}$ and $\hat{x} \otimes y \otimes \hat{z}$ must be positive. These two terms cannot be reduced to a single rank-1 term because the span in the two first components has dimension two. Adding any additional nonzero terms from R_1 cannot reduce the number of orthogonal decomposed tensors in the sum.

So, if A_1 is the first term, we cannot express A as the sum of fewer than three decomposed tensors.

6. Conclusion. We conclude that the Eckart-Young approximation theorem cannot be extended to the orthogonal rank tensor decomposition. In §3, we considered the orthogonal rank decomposition of A given by

$$A = \sigma_1 U_1 + \sigma_2 U_2.$$

If we can extend the Eckart-Young approximation theorem, then $\sigma_1 U_1$ should be the best rank-1 approximation, but we saw in §4 that this is not the case. On the other hand, the orthogonal rank decomposition is not unique [4], so in §5 we considered the alternate tack of extending A_1, the best rank-1 approximation, to an orthogonal rank decomposition. In this case, we found that we cannot express A using fewer than three terms whenever A_1 is the first term.

In other words, the best rank-1 decomposition is not nested in the best rank-2 decomposition. Thus we have derived a counterexample to the extension of the Eckart-Young matrix approximation theorem to the orthogonal rank tensor decomposition.

Acknowledgments & Comments. I am very grateful to Didier Leibovici for email exchanges that inspired the present work and for bringing to my attention related work on the Higher-Order SVD (HOSVD) by De Lathauwer, De Moor, and Vandewalle [1, 2]. I would also like to acknowledge the anonymous referees for their extremely helpful critiques and recommendations and to thank Nick Higham for his handling of this manuscript.

REFERENCES

[1] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Analysis, 21 (2000), pp. 1253–1278.

[2] ———, *On the best rank-1 and rank-($R_1, R_2, \ldots, R_n$) approximation of higher-order tensors*, SIAM J. Matrix Analysis, 21 (2000), pp. 1324–1342.

[3] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.

[4] T. G. Kolda, *Orthogonal tensor decompositions*, SIAM J. Matrix Analysis, 23 (2001), pp. 243–255.

[5] D. Leibovici and R. Sabatier, *A singular value decomposition of a k-way array for principal component analysis of multiway data, PTA-k*, Linear Algebra Appl., 269 (1998), pp. 307–329.